Valuation in the US Commercial Real Estate*

Eric Ghysels † **Alberto Plazzi** ‡ **Rossen Valkanov** §

This Version: October 2006 ¶

*This paper is a direct result from a presentation at the EFMA, Madrid, 2006. The authors thank participants at that conference for stimulating discussions and Jun Liu and Allan Timmermann for useful comments. All remaining errors are our own.

†Department of Economics, University of North Carolina, Gardner Hall CB 3305, Chapel Hill, NC 27599-3305, phone: (919) 966-5325, e-mail: eghysels@unc.edu.

‡The Anderson School at UCLA and SAFE Center University of Verona, 110 Westwood Plaza, Los Angeles, CA 90095-1481, phone: (310) 825-8160, fax: (310) 206-5455, e-mail: alberto.plazzi.2010@anderson.ucla.edu.

§Corresponding author. Rady School at UCSD, Pepper Canyon Hall, 9500 Gilman Drive, MC 0093, La Jolla, CA 92093, phone: (858) 534-0898, fax: (858) 534-0745, e-mail: rvalkanov@ucsd.edu.

¶The latest draft is available at: http://www.rady.ucsd.edu/valkanov

Abstract

We consider a log-linearized version of a discounted rents model to price commercial real estate as an alternative to traditional hedonic models. First, we verify a key implication of the model, namely, that cap rates forecast commercial real estate returns. We do this using two different methodologies: time series regressions of 21 US metropolitan areas and mixed data sampling (MIDAS) regressions with aggregate REITs returns. Both approaches confirm that the cap rate is related to fluctuations in future returns. We also investigate the provenance of the predictability. Based on the model, we decompose fluctuations in the cap rate into three parts: (i) local state variables (demographic and local economic variables); (ii) growth in rents; and (iii) an orthogonal part. About 30% of the fluctuation in the cap rate is explained by the local state variables and the growth in rents. We use the cap rate decomposition into our predictive regression and find a positive relation between fluctuations in economic conditions and future returns. However, a larger and significant part of the cap rate predictability is due the orthogonal part, which is unrelated to fundamentals. This implies that economic conditions, which are also used in hedonic pricing of real estate, cannot fully account for future movements in returns. We conclude that commercial real estate prices, at least at an aggregate level, are better modeled as financial assets and that the discounted rent model might be more suitable than traditional hedonic models, at least at an aggregate level.

1 Introduction

It is often argued that real estate is unlike other financial assets. It is perhaps because of that belief that the pricing of properties is approached quite differently from the pricing of other financial assets. Indeed, the prevalent method for valuing real estate, based on the work of Rosen (1974) and Rosen and Topel (1988), is to construct a hedonic price index of a property with given characteristics (see also Poterba (1991), DiPasquale and Wheaton (1994), and Mayer and Somerville (2000)). The alternative of looking at a real estate property as any other financial asset and pricing it with the present discounted value of expected rents has received almost no consideration in the academic literature.

In this paper, we use a log-linearized version of the discounted rent model to price real estate assets. We argue that this approach is particularly suitable for valuing commercial properties (office building, apartment, retail, or industrial space) which are the main focus of the paper. Following Plazzi, Torous, and Valkanov (2006), we price a commercial property as the present value of its expected net rents.[1] Since it is well-known in the commercial and residential real estate literature that the value of a property in a given metropolitan area is a function of demographic, local economic, and geographic determinants (Capozza, Hendershott, Mack, and Mayer (2002), Abraham and Hendershott (1996), Lamont and Stein (1999), Malizia (1991), Rosen and Topel (1988), among others), we let the expected returns and the expected growth in rents in the model to depend on local state variables. This modeling approach parsimoniously captures the observed time-variation and regional cross-sectional differences of real estate valuations.

A direct implication of our approach is that the cap rate, i.e., the rent-to-price ratio, is related to future commercial real estate returns. We test this predictive relation using a unique dataset of market-based cap rates and commercial real estate returns for 21 metropolitan areas in the US over the 1985 – 2002 sample. The data is provided by Global Real Analytics (GRA) and is available at bi-annual frequency. More specifically, we estimate a forecasting regression of future long-horizon returns on lagged cap rates and find that in 17 (14) of the 21 regions the cap rate predicts yearly returns at the 10% (5%) level. We show that the results are not only statistically but also economically significant. The significance of the predictability results is quite encouraging, especially given the small sample size.

A concern with real estate predictability tests is that they might be driven by a

[1] Net rents is defined as rents minus any operating expenses adjusted for vacancies.

mechanical correlation between cap rates and future returns. This issue arises if real estate prices do not truly reflect market valuations (because of frictions), but are to some extent influenced by appraisals. Indeed, since cap rates are often used to formulate appraisals which in turn have an effect on market prices, then the artificial link between cap rate and future returns would obtain. To investigate this possibility, we use real estate investment trust (REIT) returns which reflect true market valuations of commercial real estate properties. REITs are traded on stock exchanges and are not influenced by appraisals. Moreover, since the REIT returns are from CRSP, which is a source completely unrelated to the GRA-provided cap rates, this provide an added level of robustness to our methods. The only technical complication is that REIT returns are available at daily frequency whereas cap rates are observed bi-annually.

We investigate the relation between the bi-annual cap rates and future daily REIT returns using a mixed data sampling framework (MIDAS). The advantage of the MIDAS approach is that the forecasting relation is estimated by taking into account all the information in the series of daily returns. We estimate two variations of MIDAS predictive regressions and find that the cap rate forecasts future daily returns at conventional levels of significance. The forecasting ability in the MIDAS regressions is comparable to that observed in the initial predictive regressions. Moreover, the economic magnitude of the forecasting relation is economically significant. These results suggest that the documented predictive relation is indeed a robust feature of our data. As an aside, the application of MIDAS to predictive regression is a novel approach. Indeed, to our knowledge, MIDAS regressions have never been used in a predictive context.

To understand the provenance of the documented predictability, we use our model to express the cap rate as a function of the same local state variables that drive expected returns and growth in rents. We show that the cap rate can be decomposed into three components: (i) local state variables (demographic and local economic variables); (ii) growth in rents; and (iii) an orthogonal part. The third component is the residual from regressing the cap rate on the first two components. Importantly, since the cap rate is a real estate valuation measure, its linear connection to local economic, demographic, and geographic variables is reminiscent of hedonic models. In fact, we implicitly show that hedonic real estate pricing is not inconsistent with our framework. On the contrary, under certain reasonable assumptions, the log-linearized version of the discounted rent model yields a version of the hedonic relation. Moreover, the empirical test of our cap rate decomposition is identical to hedonic regressions. Regressing the cap rate on local state variable and growth in rents, we find that the regressors account for about 30 percent of the time-series variation in the valuation measure. In retrospect, the success of our variables to capture fluctuations in the cap rate should not come as a surprise given the wealth of evidence in support of hedonic models.

We have thus far argued that the cap rate forecasts future commercial real estate returns and that local state variables explain a significant fraction of the fluctuations in the cap rate. The natural next step is to investigate whether the observed predictability can be traced to the economic variables. If this is indeed the case, then hedonic variables and the log-linearized version of the discounted rent model will be equally successful at predicting movement in future returns. Based on the previous results, we write the realized cap rate as the sum of expected cap rate and an orthogonal part. The expected cap rate is a linear combination of the local state variables whereas the orthogonal part is the portion that cannot be captured by these variables. We regress future commercial real estate returns on both components of the cap rate. The economic variables do have some forecasting power. In 15 (14) out of the 21 regions, the expected cap rate explains movements in returns at the 10% (5%) level. More interestingly, a large part of the predictive ability of the cap rate is due to the orthogonal part that cannot be explained by the local state variables. Indeed, in 11 (8) out of the 21 regions, the orthogonal part explains movement in future returns at the 10% (5%) level. This finding is corroborated by additional statistical tests. In sum, the economic variables that are also used in hedonic pricing models cannot fully account for the future movement in prices. We conclude that commercial real estate return in our dataset are better modeled as financial assets and that the log-linearized version of the discounted rent model is more suitable than hedonic model, at least at an aggregate level.

The plan of the paper is as follows. In Section 2, we present our valuation framework and discuss its application to commercial real estate. We also connect our model to the traditional hedonic real estate models. In Section 3, we discuss the commercial real estate data. The main predictive results are presented in Section 4. We also investigate the provenance of the predictability and attempt to reconcile the results with hedonic regressions. In section 5, we use MIDAS predictive regressions as an alternative approach to document the predictability. We offer concluding remarks in Section 6.

2 The Model: Commercial Real Estate Valuation

The gross return of a commercial property (say, an apartment building) in metropolitan area i (say, San Diego, California) from t to $t+1$ can be defined as $1+R_{i,t+1} \equiv (P_{i,t+1}+H_{i,t+1})/(P_{i,t})$, where $P_{i,t}$ is the price of the property at the end of period t and $H_{i,t+1}$ are the net rents (i.e., rent minus any operating expenses adjusted for vacancies) from period t to $t+1$. This definition of return is similar to that of any other asset, just considers the fact that commercial properties provide real estate services at a market price $H_{i,t+1}$.

If we take a log transformation of the return definition, then we can write the log return, $r_{i,t+1} \equiv \log(1+R_{i,t+1})$, of a commercial property as $r_{i,t+1} \approx \kappa_i + \rho_i p_{i,t+1} + (1-\rho_i)h_{i,t+1} - p_{i,t}$, where $p_{i,t} \equiv \log(P_{i,t})$ is the log price and $h_{i,t+1} \equiv \log(H_{i,t+1})$ is log net rent. This expression is obtained, following Campbell and Shiller (1988), from a first-order Taylor approximation to the log return expression. The constants κ_i and ρ_i are derived from the linearization.[2] Solving this relation forward, imposing the transversality condition $\lim_{k\to\infty} \rho_i^k p_{i,t+k} = 0$ to avoid the presence of rational bubbles, and taking expectations at time t, gives the following present value relation for the log price $p_{i,t}$ of a commercial real estate property in area i:

$$p_{i,t} = \frac{\kappa_i}{1-\rho_i} + E_t\left[\sum_{k=0}^{\infty} \rho_i^k \left[(1-\rho_i)h_{i,t+1+k} - r_{i,t+1+k}\right]\right] \tag{1}$$

The pricing relation (1) expresses the value of a commercial property in terms of expected cash flows (net rents) and discount rates. A high property price today reflects the expectation of high future rents or of lower future expected returns or both. If commercial real estate markets are efficient, then information about future cash flows or future discount rates should be reflected in current property prices.[3]

The rent-price ratio, $H_{i,t}/P_{i,t}$, is known as the "cap rate" in the real estate literature (Geltner and Miller (2000)). If we define the log cap rate as $cap_{i,t} \equiv h_{i,t} - p_{i,t}$, then from expression (1) we can write

$$cap_{i,t} = -\frac{\kappa_i}{1-\rho_i} + E_t\left[\sum_{k=0}^{\infty} \rho_i^k r_{i,t+1+k}\right] - E_t\left[\sum_{k=0}^{\infty} \rho_i^k \Delta h_{i,t+1+k}\right] \tag{2}$$

[2]See Campbell, Lo, and MacKinlay (1997) for more details on the log-linearization. In brief, $\rho_i \equiv 1/(1+exp(\overline{h_i - p_i}))$, being $\overline{h_i - p_i}$ the average log rent - price ratio in area i.

[3]Expression (1) has been previously used in the asset pricing literature to analyze the fluctuations of equity returns (see Campbell and Shiller (1988) and Campbell (2003) for a review).

under the condition that expected returns and expected growth in rents are stationary.

The above expression obtains from re-writing the pricing equation (1) in terms of stationary and co-integrated terms. Expression (2) is easiest to understand as a consistency relation. It states that if a cap rate is high, then either the property's expected return is high, or the expected rental growth is low, or both. A key feature is that the cap rate is a state variable that is readily observable in the market and embodies all relevant information about future expected returns and rent growth.

The implication of (2) that $cap_{i,t}$ is a possible forecaster of future returns has given rise to a large return predictability literature, notably with the real estate investment trust (REIT) returns and cap rates (Karolyi and Sanders (1998), Ling, Naranjo, and Ryngaert (2000), Liu and Mei (1992), and Nelling and Gyourko (2000), among others).

2.1 The Cap Rate and Future Commercial Real Estate Returns

As a result of our pricing equation and to the extent that future realized returns proxies for expected returns, the cap rate can be used to explore fluctuations in both these variables. With real estate assets, this relation is more likely to hold at horizons of one year or more, when the short-horizon frictions that exist in these markets become less of a concern. In the context of commercial real estate, the forecasting ability of the cap rate in predicting future excess returns has been documented by Plazzi, Torous, and Valkanov (2006). They work on a larger cross-section of areas spanning a shorter period and focus on the dynamics within each property type. To this extent, they pool observations across metropolitan areas to improve efficiency of the estimates and rely on a double resampling procedure. Our approach is closely related to their work, although with a different point of view: we are mainly interested in capturing pricing differences across metropolitan areas rather than measuring differences in cap rate predictability across property types (we discuss this issue further in Section 4).

Based on the above expressions, one can test the forecasting ability of the cap rate in a given area by running the following regression:

$$r_{i,t+1\rightarrow t+k} = \alpha_{i,k} + \beta_{i,k}\left(cap_{i,t}\right) + \varepsilon_{i,t+k} \tag{3}$$

Relation (2) predicts that the (log) cap rate should be positively related to future excess returns (in excess of the Tbill rate) in each metropolitan area. The long-horizon nature of the predictive relation is suggested by expression (2) and the horizon of the returns is denoted by k. Whether or not this relation is economically and statistically significant over suitable long horizons (in our case, one year, or $k = 2$) is ultimately an empirical question which we address in the next sections.

The forecasting regression (3) parallels the literature on predictability of stock returns. Despite the similarity of the cap rate as valuation ratio with the dividend-price or earnings-price ratios used in the equity literature, the pricing analogy between the equity and real estate markets should be used with caution. For instance, it is well known that real estate prices in a metropolitan area are very sensitive to local economic conditions, demographic trends, and geographic location, much more so than are prices of other assets. Therefore, in order to fully characterize the results of our predictive regressions, we need to take into account the impact of these underlying local state variables on the cap rate.

2.2 The Cap Rate and Local State Variables

To proceed further, we need to make explicit assumptions about the form of expected returns, $E_t r_{i,t+1}$, and the expected rental growth rates, $E_t \Delta h_{i,t+1}$. It is well-known in the commercial and residential real estate literature that the pricing of properties across metropolitan areas is a function of demographic, local economic, and geographic determinants. For instance, Capozza, Hendershott, Mack, and Mayer (2002) find that house price dynamics vary with city size, income growth, population growth, and construction costs. Abraham and Hendershott (1996) document a significant difference in the time-series properties of house prices in coastal *versus* inland cities. Lamont and Stein (1999) show that house prices react more to city-specific shocks, such as shocks to per-capita income, in regions where homeowners are more leveraged. Miles, Cole, and Guilkey (1990) build a transaction-based index based on a pricing model where commercial real estate returns depend on local market economic health measures (population, total employment, unemployment rate, total per capita personal income and per capita personal income by employment sector for the demand side and construction and finance insurance and real estate earnings for the supply side), location measures and physical structure measures. Malizia (1991) analyzes metro-level employment, income and population forecasts used by real estate analysts, investors and developers in order to estimate anticipated absorption for the proposed project's market area. Capozza, Hendershott, Mack, and Mayer (2002) analyze the effect on price dynamics of demand and supply variables affecting transaction frequencies. They find that city size, income growth, population growth, and construction costs can explain differences in serial correlation and mean reversion of housing prices across metro areas.[4] Finally, the entire hedonic real estate pricing literature (Rosen and Topel (1988), Poterba (1991), DiPasquale and Wheaton (1994), and Mayer and Somerville (2000)) takes into account these factors.

In light of this evidence, we model the expected return in metropolitan area i as

$$E_t r_{i,t+1} = r_i + \delta_i x_{i,t} \tag{4}$$

where $x_{i,t}$ is a vector of demographic, local economic, and geographic variables that capture differences across metropolitan areas and r_i is the unconditional expected return for that area. This specification implies that the variables in $x_{i,t}$ capture risk factors for the specific area.

The growth in rents in metropolitan area i can also be expressed as

$$E_t \Delta h_{i,t+1} \quad = \quad g_i + \tau_i x_{i,t} + y_{i,t} \tag{5}$$

where g_i is the unconditional expected rental growth rate in the area and $y_{i,t}$ is the variation in rent growth that is orthogonal to the variation in expected returns. These are variations of cash flows that are not compensated by a systematic increase in risk. Specification (5) allows rent growth and expected returns in area i to be correlated. For example, if $\tau_i = 1$ then both rental growth and expected returns respond equivalently to changing economic conditions.[5]

The metropolitan state variables $x_{i,t}$ and $y_{i,t}$ are likely to be persistent over time. We capture this time dependence by allowing both processes to follow autoregressive $AR(1)$ processes, or $x_{i,t} = \phi_i x_{i,t-1} + \xi_{i,t}$ and $y_{i,t} = \psi_i y_{i,t-1} + \zeta_{i,t}$, where $\xi_{i,t}$ and $\zeta_{i,t}$ are uncorrelated contemporaneously at all leads and lags.

Substituting the expected returns and expected rent growth expressions, (4) and (5), in the cap rate expression (2) and solving forward yields the following expression for the cap

[4] While most of the cited papers focus strictly on the residential market, similar mechanisms are likely at play in commercial real estate.

[5] Lettau and Ludvigson (2005) use a similar specification to model the correlation between expected returns and expected dividend growth in common stocks.

rate (omitting the κ term)

$$cap_{i,t} = \left(\frac{r_i - g_i}{1-\rho_i}\right) + \left(\frac{\delta_i(1-\tau_i)}{1-\rho_i\phi_i}\right) x_{i,t} - \left(\frac{1}{1-\rho_i\psi_i}\right) y_{i,t}. \tag{6}$$

In equation (6), fluctuations in the cap rate must be captured by state variables in $x_{i,t}$ and the determinants of rent that are orthogonal to expected returns, $y_{i,t}$. If we were interested in the structural parameters r_i, g_i, δ_i, τ_i, ρ_i, ϕ_i, and ψ_i, then we could estimate them in a number of ways. For instance, if $y_{i,t}$ is observable, then a GMM estimation will be straightforward. Alternatively, if $y_{i,t}$ is unobservable and there are no good proxies for it, then a Kalman filtering approach could be applied under certain distributional assumptions of $\xi_{i,t}$, $\zeta_{i,t}$, and $\zeta_{i,t}$. However, we are not interested in the structural parameters. Our primary goal is to decompose fluctuations in the cap rate into two parts: an expected component, determined by movements in all the local state variables $x_{i,t}$ and $y_{i,t}$, and a residual part that cannot be explained by these variables.

We re-write equation (6) in the following semi-structural form:

$$cap_{i,t} = \mu_i + \lambda_i^x x_{i,t} + \lambda_i^y y_{i,t} + \upsilon_{i,t} \tag{7}$$

where $\mu_i = \left(\frac{r_i - g_i}{1-\rho_i}\right)$, $\lambda_i^x = \frac{\delta_i(1-\tau_i)}{1-\rho_i\phi_i}$, and $\lambda_i^y = -\frac{1}{1-\rho_i\psi_i}$. The parameters μ_i, λ_i^x, and λ_i^y in the above expression can be estimated with a simple regression for each MSA.[6] This is a semi-structural relation, because it is derived from a structural expression, where not all parameters are identifiable.

In order to investigate whether fluctuations in the cap rate are due to variations in fundamentals, we regress $cap_{i,t}$ on $x_{i,t}$ and $y_{i,t}$. The variables in $x_{i,t}$ are extracted from demographic and local economic variables such as population employment, income and construction costs[7]. The variations in the expected growth in rents that are orthogonal to expected returns, $y_{i,t}$, are not directly observable. However, they can be identified using expression (5). Indeed, the variables in $x_{i,t}$ are observable and so is the growth of rents $\Delta h_{i,t+1}$ (as a proxy for $E_t \Delta h_{i,t+1}$). We can identify $y_{i,t}$ as the residual from a regression of rent growth rates on the local economic variables in $x_{i,t}$. Then, by construction, $y_{i,t}$ will be

[6]In this case, the OLS and GMM estimates will be identical, since the systems of equations is just -identified.

[7]Since we are not pooling the metropolitan areas, we are unfortunately not able to capture any coastal effect which is by its nature constant over time.

the variation in $\Delta h_{i,t+1}$ that is orthogonal to $x_{i,t}$.

Since the variables in $x_{i,t}$ and $y_{i,t}$ are now available, we estimate expression (7) with least squares. The residuals from this regression represent the part of the cap rate (or the valuation measure of commercial real estate) that is not captured by fundamentals, i.e., expected returns or future cash flows. Hence, the residual $\upsilon_{i,t}$ can be interpreted as the mispricing in the valuation of the property in metropolitan area i at time t.

2.3 A Decomposition of the Relation between the Cap Rate and Future Commercial Real Estate Returns

Once we have related the cap rate to fundamentals, we are now able to fully characterize the forecasting regression results from expression (3) and disentangle the forecasting ability of the cap rate into three components. The first one reflects changes in local economic conditions, proxied by the variables in $x_{i,t}$. The second one captures rent growth components orthogonal to these variables, $y_{i,t}$, derived from expression (5). Finally, the last component, $\upsilon_{i,t}$, reflects fluctuations in the cap rate that are unrelated to fundamentals.

Substituting the expression of the cap rate (7) into (3), we obtain:

$$\begin{aligned} r_{i,t+1} &= \mu + \gamma_i^x \; x_{i,t} + \gamma_i^y \; y_{i,t} + \gamma_i^\upsilon \; \upsilon_{i,t} + \varepsilon_{i,t+k} \\ &= E\left[cap_{i,t}\right] + \gamma_i^\upsilon \; \upsilon_{i,t} + \varepsilon_{i,t+k} \end{aligned} \tag{8}$$

The variable $\upsilon_{i,t}$ succinctly summarizes all information that can be to some extent regarded as "irrational", or in other words not related to fundamentals but is still relevant for predicting returns. With expression (8), we can directly test the economic and statistical impact of this component and its relative importance in our predictive regressions. This will allow us to characterize the nature of the cap rate's forecasting ability, namely, whether its predictive power is attributable to variation in fundamental information or to unrelated factors captured by $\upsilon_{i,t}$.

Intuitively, we expect the mispricing to be less severe for areas where the economic risk factors account for more of the variation in cap rates. However, whether or not the unexpected component has a significant impact on future returns is an open issue which we empirical test in Section 4.

2.4 Hedonic Models and Our Approach

Modeling expected returns and growth in rents as a function of demographic, geographic, and local economic factors provides a natural opportunity to link our method to the hedonic pricing literature. In hedonic regressions, these factors are explicitly taken into account, along with property-specific characteristics. While the specifications in expressions (7) and (8) are similar to the hedonic models, there are several important differences in our implementation.

The differences are due to the data as much as the approach. For instance, our data consists of portfolios of real estate properties in a given region rather than individual properties. Unfortunately, the data-provider, Global Real Analytics (GRA), was unable to give us access to the disaggregated data. Hence, we cannot take into account property-specific features that some hedonic models would include in the explanatory variables. Our analysis could be viewed as capturing the behavior of the "average property" in a given area. In constructing these portfolios, GRA has made every effort to hold quality constant (see data section below). For instance, for apartments, we have a pool of property A (luxury) apartments for each MSA. The implicit assumption is that the quality of the commercial property and its other characteristics are accurately taken into account when constructing the portfolios.

Working with portfolios comes with advantages and drawbacks. On the positive side, in portfolio returns the idiosyncratic noise in the property data is largely attenuated. A disadvantage of the portfolio approach is that we don't know whether the pool mix of type A apartments in some metropolitan areas has the same characteristics (bedrooms, bathrooms, etc.) as the same properties in another area. While there are always omitted controls in such aggregations, we expect the cross-sectional difference in pricing to be relatively small. Moreover, as long as this cross-sectional heterogeneity does not change over time, our time-series result should not be affected by it.

3 The Commercial Real Estate Data

3.1 Returns, Cap Rates, and Growth in Rents

In the commercial real estate data, we have prices and annualized cap rates of class A offices, apartments, retail and industrial properties for twenty one U.S. metropolitan areas. The data are provided by Global Real Analytics (GRA) and are available on a semi-annual basis beginning with the second half of 1985 (1985:2) and ending with the second half of 2002 (2002:2). We list these metropolitan areas in Table A1 of the Appendix. The prices and cap rates for each property category are averages of transactions data in given six months. Taken together, we have a panel of 140 observations (35 time-series $\times$ 4 property types) for each metropolitan area. This data is also used by many real estate, financial, and government institutions[8]. We consider this as an indication, albeit not scientifically rigorous, of the data's accuracy.

Given annual cap rates, CAP_t, and prices, P_t, of a particular property type in a given area, we construct semester t's net rents as $H_t = (\mathrm{CAP}_t \times P_t)/4$. The gross returns at $t+1$ are then obtained as $1 + R_{i,t+1} = (P_{i,t+1} + H_{i,t+1})/(P_{i,t})$, while H_t/H_{t-1} gives one plus the rent growth. For consistency with the previously derived expressions, we work with log cap rates, $cap_t = \ln(CAP_t)$, and log rental growth rates, $\Delta h_t = \ln(H_t/H_{t-1})$. Also, we rely on log excess returns, $r_t = \ln(1+R_t) - \ln(1+R_t^{Tbl})$, where R_t^{Tbl} is the three month Treasury bill yield. Table A1 in the Appendix also reports time-series averages of excess returns, rental growth rates, and cap rates for all property types across all metropolitan areas.

3.2 Demographic and Local Economic State Variables

To account for differences across metropolitan areas we use the following control variables: population growth ($gpop_{i,t}$), the growth of income per capita ($ginc_{i,t}$), and the growth of employment ($gemp_{i,t}$), all of which are provided by the Bureau of Economic Analysis at an annual frequency. We also use the annual growth in construction costs ($gcc_{i,t}$) compiled by R.S. Means. The construction cost indices include material costs, installation

[8]A partial list of the subscribers includes Citigroup, GE Capital, J.P. Morgan/Chase, Merrill Lynch, Lehman Brothers, Morgan Stanley Dean Witter, NAREIT, Pricewaterhouse-Coopers, Standard & Poors, Trammell Crow, Prudential RREEF Funds Capital/Real Estate Investors, Washington Mutual, FDIC, CalPERS, and GMAC.

costs, and a weighted average for total in place costs. In addition, after lagging by two years, we include log population ($pop_{i,t-2}$), log per capita income ($inc_{i,t-2}$), log employment ($emp_{i,t-2}$), and log construction costs ($cc_{i,t-2}$), to proxy for the level of urbanization (Glaeser, Gyourko, and Saks (2004)). We lag these level variables by two years to prevent a mechanical correlation with corresponding growth rates. All these variables are available for each metropolitan area at annual frequency. Since our real estate data come at biannual frequency, we assume these variables to be constant through the year[9].

3.3 REIT Data

We use real estate investment trust (REIT) returns as an additional source of commercial real estate data. The REIT portfolio return is the CRSP value-weighted REITs index, available at daily frequency for the 1985-2004 period. This index combines stock price and returns data on all REITs that have traded on the NYSE, AMEX and NASDAQ exchanges during the sample period. We use the REITs as an important robustness check of our results in section 5.

4 Results

4.1 The Predictability of Commercial Real Estate Returns

We first present the results from the predictability expression (3). For each metropolitan area, we run a time-series regression of future one-year ($k = 2$) non-overlapping returns on lagged cap rate for the entire 1986 to 2002 period. Ideally, we would have liked to run this regression for every property type in a given metropolitan area. However, we have a small time-series for each property type and area and, consequently, the statistical power of our forecastability tests would be low. Therefore, we pool the observations for all four property types and run one regression for each MSA. This approach is reasonable for two main reasons. First, our control variables are available at the metropolitan area level and this information is of fundamental importance in understanding pricing differences. Second, we are mainly interested in identifying which areas exhibit symptoms of mispricing that, we claim, are of higher order of importance than those between property types of the same area. We account for cross-correlation between different property types by computing the

[9]This will therefore reduce the power of our test, as there will be less variation in our control variables.

t-statistics using Newey-West standard errors with 4 lags. Finally, it is worth mentioning that the use of non-overlapping returns renders our statistical results less prone to the issue of spurious correlation in residuals induced by evaluating overlapping returns, which make reliable statistical inference harder to obtain (Valkanov (2003)).

The estimates and t-statistics from these regressions are presented in Table 1 for each metropolitan area along with the associated R^2s. All coefficient estimates are positive, as expected from the model in equation (2). Out of the 21 regions, 16 are significant at the 5 percent level (or about four-fifths of the regions). The non-significant regions are Charlotte (North Carolina), Denver (Colorado), Houston (Texas), San Francisco (California) and Seattle (Washington). The fact that these regions are not geographically and demographically close to each other suggests that the insignificant results are unlikely to be driven by a common cross-sectional factor.

To gauge the economic significance of our results, we calculate the effect of a two-standard deviation shock to a region's lagged (log) cap rate on the next year returns using the coefficient estimates from column $\widehat{\beta}$ of Table 1. This statistic, reported in column $ec_{\hat{\beta}}$ of Table 1, is in percents. We also report a two-standard-deviation confidence interval for this value by using the standard errors of the coefficient $\widehat{\beta}$. The lower and upper bounds of this confidence interval are denoted by $ec^{-}_{\hat{\beta}}$ and $ec^{+}_{\hat{\beta}}$, respectively. Finally, we report the absolute value of this magnitude as a fraction of market return volatility in the last column of the table. For instance, the estimated coefficient for the cap rate in the case of Atlanta is 0.248 and plus or minus two-standard errors yields 0.009 and 0.488 as the lower and upper bounds, respectively. A two standard deviation shock in the biannual log cap rate of Alabama is (2*6.34%). This leads to an expected change in next years market return of 3.1% (0.248*2*6.34%). The lower and upper bounds on this estimate are obtained by using 0.009 and 0.488 instead of 0.248. We observe that a two standard deviation shock explains 44 percent of the return volatility which is very significant in economic terms. Moreover, similar regressions in the stock market (returns on lagged dividend yield) yield much lower estimates partly because the estimates of the predictability parameter is lower and also because the standard deviation of the dividend yield is low.

The last four columns of Table 1 suggest that the cap rate is not only statistically but also economically significant predictor of commercial real estate returns in most MSAs. Understanding the reasons for this predictability and, in particular, whether it can be due to underlying economic fluctuations, is a questions we aim to address. Before doing so, we have to understand what drives the time-series fluctuations of the predictor, the cap rate. This is the topic of the next section.

4.2 Cap Rate Decomposition

In this section, we decompose the cap rate into fluctuations due to local state variables ($x_{i,t}$), growth in rents orthogonal to economic fluctuations ($y_{i,t}$), and orthogonal parts ($\upsilon_{i,t}$) using regression (7). In order to do that empirically, we first have to identify the variables $x_{i,t}$ and $y_{i,t}$.

We construct $x_{i,t}$ from the eight local economic and demographic variables described in section 3.2. While we could directly use the raw variables in our regressions, some of them are highly correlated. This is not surprising as these variables all proxy for economic conditions and thus for the risk factors affecting prices and rents dynamics. The high correlation results in low t-statistics which implies that multicollinearity is an issue. In order to reduce the number of correlated variables in the regression and to parsimoniously summarize their information, we perform a principal component analysis (PCA, henceforth) of the (standardized) economic variables, at each point in time. Our analysis reveals that four out of the eight principal components are particularly correlated with the level and growth in population and income as well as with the level of construction costs, and account for more than 80% of the overall volatility. The four extracted principal components (which, by construction, are orthogonal) constitute our $x_{i,t}$ variable.

Unlike $x_{i,t}$, the part of rent growth that is not explained by expected returns, $y_{i,t}$, is not directly observable. To identify $y_{i,t}$, we use equation (5) and regress the bi-annual growth in rents of each metropolitan area on lagged $x_{i,t}$.[10] The residuals from this regression are, by construction, the part of growth in rents that is orthogonal to the economic principal components in $x_{i,t}$.

We next estimate the semi-structural equation (7) by regressing bi-annual log cap rates on $x_{i,t}$ and $y_{i,t}$ for each metropolitan area. This is the cap rate decomposition regression the results of which are displayed in Table 3. The Table contains the coefficient estimates of the four economic factors $x_{i,t}$ and of the $y_{i,t}$ as well as their Newey-West t-statistics and the R^2s. The four components in $x_{i,t}$ all have good explanatory power. For instance, the first

[10] The results from this regression are not presented for conciseness but are available upon request.

three components are significant at the 5% level in 16 out of the 21 regressions, and the fourth is significant at the same level in 10 out of the 21 regressions.[11] By contrast, in only one metropolitan area, Chicago, is the coefficient in front of $y_{i,t}$ significant at the 5% level. Since $y_{i,t}$ is the part of future rent growth that is orthogonal to expected returns, this result is consistent with Campbell (1991)'s claim in the equity literature that the state variable (dividend yield in their case) is not correlated with future growth in cash flow (growth in rents in our case and growth in dividends in the equity literature). Overall, the goodness of fit in these regressions is surprisingly good, especially if we consider the modest sample size. The average and median R^2s are 0.306 and 0.279, respectively. The lowest R^2 of 0.113 is observed in the Dallas-Forth Worth area. In two regions, Riverside-San Bernardino and San Diego, the goodness of fit is as high as 0.673 and 0.620, respectively.

4.3 Understanding the Predictability Results

In this section, we investigate the provenance of the predictability documented in section (4.1). Understanding the economic reasons for the results is important because if the predictability is due to fundamental fluctuations in $x_{i,t}$ and $y_{i,t}$, then our approach will be empirically indistinguishable from the hedonic models, because expected returns and growth in rents are a function of the local economic variables. Alternatively, if the predictability of the cap rate is not due to these variables, then our approach would dominate the hedonic models.

In table 3, we present the results from regression (8). As explained in section (2.3), this regression represents a decomposition the cap rate predictability documented in Table 1 into three variables: $x_{i,t}$, $y_{i,t}$, and $\upsilon_{i,t}$. A priori, at least one of these variables ought to predict returns, because taken together they account for the entire fluctuation in the cap rate (see regression (7)). T-statistics are reported below all estimates along with R^2s for each metropolitan area.

The four local economic components in $x_{i,t}$ are statistically significant for most of the MSAs. In fact, for most regions, more than one economic variable is statistically significant at conventional levels. For instance, for Atlanta, three out of the four economic variables are significant at the 10 percent level and two are significant at 5 percent. For Baltimore, one $x_{i,t}$ factor is significant at the 5 percent level, and so on. Only for three out of the 21 regions,

[11] By random chance, we expect about one significant result (or 0.05*21).

Riverside-San Bernardino, San Francisco, and Tampa/St. Petersburg, no economic variable in $x_{i,t}$ predicts future returns. Overall, this suggests that the factors that we have chosen have some forecasting power for future commercial real estate returns. This evidence also confirms that if these factors are part of a hedonic model, they will have some explanatory power.

The $y_{i,t}$ variable contributes only modestly toward the predictability results. In the table, the coefficients in front of only 6 (4) regions are significant at the 10 (5) percent level. Interestingly, one of the largest and most statistically significant coefficient obtains in the Tampa/St. Petersburg regression. This is the region for which the economic variables $x_{i,t}$ failed to explain the predictability. Hence, there might be some interesting complementarity between fluctuations in expected returns and growth in rents.

The variable $\upsilon_{i,t}$ captures fluctuations in the cap rate that are orthogonal to $x_{i,t}$ and $y_{i,t}$. These fluctuations cannot be explained by underlying economic factors. In table 3, we observe that the coefficients in front of the $\upsilon_{i,t}$ variable are significant in 10 (8) out of the 21 regions at the 10 (5) percent level. Moreover, the sign of the estimates for all MSAs is positive. This striking evidence suggests a large fraction of the predictability cannot be accounted for by $x_{i,t}$ and $y_{i,t}$. To the extent that these variables are used in hedonic models, it also implies that the log-linearized discounted rents model and the cap rate might do a better job at explaining the future movement in commercial real estate returns.

As an alternative way of testing which variables contribute toward the predictability, we investigate whether the coefficients in regression (8) are correlated with the coefficients in regression (3) across regions. As long as the variables $x_{i,t}$, $y_{i,t}$, and $\upsilon_{i,t}$ capture the predictability, their coefficients must be positively correlated with the coefficients of the cap rate across regions. In figure 1, the six scatter plots display the relation between the four sets of coefficients in front of $x_{i,t}$, the coefficients in front of $y_{i,t}$, and the coefficients in front of $\upsilon_{i,t}$ and the coefficient in front of the cap rate. In each plot, we also display the OLS line from regressing one set of coefficients on another cross-sectionally. Finally, we also report in the plots the slope coefficient from the regressions, the t-statistics, and the R^2s.

In figure 1, the coefficient in front of $\upsilon_{i,t}$ have the strongest correlation with the cap rate coefficients. The lower-right plot which displays this correlation has the best fit. The $t-$statistic is statistically significant and the R^2 is quite large, which is quite encouraging given that we only have 21 observations in the regression. In the four $x_{i,t}$ plots, the relation with the cap rate coefficient is positive but insignificant. The $y_{i,t}$ plot is negative, as predicted by the model, but also statistically insignificant. Hence, to the extent that $x_{i,t}$ and $y_{i,t}$ are also used in hedonic models, these models will have a difficult time to predict commercial real estate returns as well as the cap rate.

It is difficult to understand why local economic variables do not explain a larger fraction of future commercial real estate returns. A potential explanation is that better proxies for $x_{i,t}$ or a better way of identifying $y_{i,t}$ might produce better results. Alternatively, we might have altogether omitted relevant information from $x_{i,t}$. In other words, we are missing some fundamental dimension of the state space. This explanation will have an effect on our results only if the neglected information is orthogonal to the current variables in $x_{i,t}$ and $y_{i,t}$ and is correlated with $v_{i,t}$. Since the list of our economic variables is quite exhaustive and is fairly standard in the real estate literature, this seems quite unlikely.

5 An Alternative Approach: MIDAS Predictive Regressions

In this section we investigate one important criticism that has been raised in predictability tests of real estate returns. It is often argued that if the prices do not truly reflect market valuations (due to frictions) but are to some extent influenced by appraisals, then a mechanical predictability relation would obtain, because cap rates are often used in appraisals (e.g., Geltner and Miller (2000) and referenced therein). In other words, cap rates are used to formulate appraisals which in turn have an effect on market prices. Hence the mechanical link between cap rate and future returns through appraisals. While this criticism is largely addressed with our transactions-based GRA data and the long-horizon regressions, there is always a small possibility that the returns with that dataset might be influenced by appraisals.

The only convincing way of addressing this concern is to turn to market returns of REITs. REIT returns reflect true market valuations and are the only real estate asset with observable transaction-based market prices. Moreover, REITs invest exclusively in commercial properties which makes them suitable for our investigation. An additional benefit is the fact that the REIT data is from CRSP, which is a source completely unrelated to the GRA-provided cap rates. Therefore, if we find a predictive relation between the cap rate and future REIT returns, the specter of "mechanical relation" will be lifted from our results.

REIT returns are available at daily frequencies whereas cap rates are observed bi-annually. This mismatch of data frequency is an opportunity and a challenge to exploit all the information in the daily data while keeping the econometric framework simple and parsimonious. We use a mixed data sampling regression (MIDAS) in order to investigate the relation between the semi-annual cap rates and future daily returns. The MIDAS approach, described below, has the advantage that it uses all the information in the series of daily REIT returns. The alternative (and often used) approach of aggregating the daily returns into semi-annual returns and then running a forecasting regression on lagged cap rate is less efficient and might obfuscate some interesting dynamics. The efficiency advantages of MIDAS regressions have been analyzed in detail by Ghysels, Santa-Clara, and Valkanov (2006), Ghysels, Sinko, and Valkanov (2006), and Ghysels and Valkanov (2006). Before presenting the results from the MIDAS predictive regressions, we briefly summarize the econometric approach.

5.1 MIDAS Methodology

A MIDAS regression, introduced by Ghysels, Santa-Clara, and Valkanov (2006) and Ghysels and Valkanov (2006) is a simple, parsimonious, and efficient way of running a regression when data is available at different frequencies. In our case, the national cap rate is available semi-annually, from the second semester of 1985 (July 1st, 1985 to December 31st, 1985) to the second semester of 2002 (July 1st, 2002 to December 30th, 2002). REITs are available daily from January 1985 to December 2004. In this section, we work with national averages rather than metropolitan areas, because REIT indices are not available for all 21 metropolitan areas. The semi-annual value-weighted average in period t of cap rates across MSAs is denoted by cap_t. The daily value-weighted return of the national REIT portfolio is denoted by $\tilde{r}_t$, where the tilde sign indicates daily returns. Also, we use the fractional differencing notation $\tilde{r}_{t-\tau}$ where τ is a fraction of a period. The length of a period in our case is six months or 130 trading days. More specifically, if t is the first period in our sample, July 1st to December 31st, 1985, then τ is an index of days during that period. A day, say, June 30th, 1985 will be denoted by $\tilde{r}_{t-1/130}$. Similarly, say, January 4th, 1986 will be denoted by $\tilde{r}_{t+4/130}$. Using this notation, we specify a MIDAS regression as

$$cap_t = \varphi + \eta \sum_{\tau=1}^{260} B(\theta;\tau)\, \tilde{r}_{t+\tau/130} + \varepsilon_t \tag{9}$$

where $B(\theta;\tau)$ is a polynomial in τ with parameters collected in the vector θ. Our main interest is whether the coefficient η, which captures the relation between cap rates and future REIT returns, is positive and statistically significant. Once the functional form of $B(\theta;\tau)$ is specified, equation (9) yields a relation between bi-annual cap rates and future daily REIT returns. The MIDAS application here is a novel approach of testing the forecasting relation suggested in equation (2). Indeed, to our knowledge, MIDAS regressions have never been used in a predictive context.

Besides the use of mixed data, regression (9) is unusual in another respect. The cap rate is now the left hand side variable (while in equation (3) it is the right-hand side variable) and future returns are on the right hand side (while in equation (3) they were on the left hand side). However, this is a legitimate regression that can best be understood as a Granger causality test as implemented by Sims (1972). Because of this reverse order, Ghysels and Valkanov (2006) call this a "reverse MIDAS" regression. The important point from their paper is that such a regression can be estimated consistently and that it captures the relation between the cap rate and future daily REIT returns.

Ghysels, Santa-Clara, and Valkanov (2006) and Ghysels, Sinko, and Valkanov (2006) discuss the importance of the functional of $B(\theta;\tau)$. The polynomial must be flexible enough but also parsimoniously parameterized. In this paper, we use two specifications for $B(\theta;\tau)$. The first one, denoted by $B_\alpha(\theta;\tau)$, places the following weight on return $\tilde{r}_{t+\tau}$

$$B_\alpha(\theta;\tau) = \frac{\exp\{\theta_1\tau + \theta_2\tau^2\}}{\sum_{j=0}^{\infty}\exp\{\theta_1 j + \theta_2 j^2\}}. \tag{10}$$

This is an exponential Almon weights parameterization, used by Ghysels, Santa-Clara, and Valkanov (2005) and others. It has several advantages. For instance, it guarantees that the weights are positive and add up to one. The functional form can produce a wide variety of shapes for different values of the two parameters. The specification is parsimonious, with only two parameters, θ_1 and θ_2, to estimate. Moreover, as long as the coefficient θ_2 is negative, the weights decay to zero as the lag length increases and the speed of the decay controls the effective number of observations used. We can increase the order of the polynomial in (10) or consider other functional forms.[12] As a practical matter, the infinite sum in (10) needs to be truncated at a finite lag. In all the results that follow, we use 260 days (which corresponds to roughly one year of trading days) as the maximum lag length[13].

The second specification that we use, denoted by $B_\beta(\tau;\theta)$, also has only two parameters in $\theta = [\theta_1;\theta_2]$ and takes the following form:

$$B_\beta(\tau;\theta) = \frac{f(\frac{\tau}{\tau^{max}}, \theta_1;\theta_2)}{\sum_{j=1}^{\tau^{max}} f(\frac{j}{\tau^{max}}, \theta_1;\theta_2)} \tag{11}$$

where $f(z,a,b) = z^{a-1}(1-z)^{b-1}/\beta(a,b)$ and $\beta(a,b)$ is based on the Gamma function, namely $\beta(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$. Specification (11) was introduced by Ghysels, Santa-Clara, and Valkanov (2006). The functional form (11) is that of a Beta distribution and we refer to it as a "Beta" polynomial from now on. Its properties are similar to those of the exponential Almon lag specification discussed above. We introduce the Beta weights as an alternative to the exponential Almon weight in order to verify that our results are not driven by a particular parametric structure.

5.2 MIDAS Results

We estimate the predictive MIDAS regression (9) with, alternatively, exponential Almon weights (10) and Beta weights (11). The estimation is carried out with non-linear least squares. The parameters φ, η, θ_1, and θ_2 in each parameterization are estimated jointly. Table 4 presents the results from these estimations. We observe that the predictability is statistically significant. The estimate of η in the Almon lag case equals 3.213 and is statistically significant. It is also significant in the Beta case, but the point estimate is slightly lower at 3.056. To compare the η in this regression with the $\widehat{\beta}$s of Table 1, we need to evaluate $1/\eta$, because of the reverse nature of our MIDAS regression. The estimate $1/\hat{\eta}$ is reported for convenience in the next column of Table 4. We observe that the $1/\eta$ coefficients of 0.311 (exponential Almon) and 0.327 (Beta) are quite similar to those in Table 1. Moreover, the goodness of fit in both regressions equal 0.354 and 0.325. These

[12]See Ghysels and Valkanov (2006) for a general discussion of the functional form of the weights.

[13]We verify that the results are not sensitive to increasing the maximum lag length beyond one year's worth of daily data.

numbers are quite large, in terms of economic significance, especially considering the small sample size.

Interestingly, the exponential Almon parameter estimates θ_1 and θ_2 are significantly different from zero. For that parameterization, when $\theta_1 = \theta_2 = 0$, the weights on future returns are equal for all 260 days. This corresponds to the case of aggregating short-horizon data into long-horizon returns. Hence, the results in table 4 imply that our forecasting MIDAS model performs better than aggregating returns and running semi-annual cap rates on semi-annual returns. We verified this claim directly by running such a regression.[14]

The weights $\hat{\theta}$ have no economic meaning and are difficult to interpret. But the polynomial $B_\alpha(\hat{\theta}; \tau)$ is of economic interest, because it capture the weights placed on future returns in the forecasting relation. We plot this polynomial for the exponential Almon case in Figure 2 with a solid line. The Beta polynomial is plotted on the same figure and same scale, with a dashed line. Both shapes are very similar, which suggests that the estimated shapes of the weights is unlikely to be due to a restrictive form of the polynomial. Moreover, the shape of the weights implies that most of the predictability occurs in the three months immediately after the cap rate is observed. For the exponential Almon weights, about 80 percent of the mass is concentrated in the first sixty days of returns. For the Beta function, about 76 percent is concentrated in the first sixty days. To summarize, the MIDAS evidence in table 4 and figure 2 confirms the findings of the previous tables.

6 Conclusion

In this paper, we use a discounted rent model as a pricing equation for commercial real estate properties. This pricing approach offers an alternative to the hedonic models that have become the de-facto norm in the real estate literature since the influential work of Rosen (1974) more than thirty years ago. We use the log-linearized version of the model to derive a connection between the cap rate and future real estate returns. Using a unique database of market cap rates and returns of commercial real estate properties in 21 U.S. metropolitan areas, we test the key implication of the model that cap rate are a good predictor of future long-horizon returns. We do that in two ways: first with simple predictive regressions, and

[14]The results of this regression are not reported here, in the interest of brevity. They are available upon request.

second, using a mixed data sampling, or MIDAS, regression and an alternative dataset of REIT returns. Both methods yield similar results. The finding that the predictability of commercial real estate returns by the cap rate is economically and statistically significant in a large fraction of the metropolitan areas validates the use of our valuation approach.

We also explore a link between our model and the hedonic literature. To do that, we allow expected returns and growth in rents to depend on local economic, demographic, and geographic variables that are also used in hedonic models. Then, using the present value relation, we express the cap rate as a function of these variables. We find that, empirically, these variables account for a significant fraction of the time-series fluctuations in cap rate. As a result, we decompose the cap rate into a projection on these variable and an orthogonal component that is not explained by the local state variables.

Using the previous cap rate decomposition, we investigate the source of the predictability and find that the economic variables account just for some fraction of it. The orthogonal component of the cap rate, that by construction is not related to changes in underlying risk factors, accounts for a statistically significant fraction of the cap rate predictability. Hence, to the extent that these economic variables are also used by hedonic models, their forecasting ability will be lower than that of the cap rate when trying to explain fluctuations of aggregate series. This would imply that the cap rate itself is a better statistics in summarizing all the relevant information in order to predict future trends.

Our results raise several interesting questions for further research. For instance, can our portfolios-based results be replicated using a more disaggregated dataset of commercial real estate properties with a larger set of hedonic characteristics? If the cap rate is not capturing fluctuations in economic activity, how can we account for its predictive ability? Is it due to mispricing captured by the cap rate? Or can the gap between the two models be bridged using additional property-specific information? We leave these questions for future research.

References

Abraham, Jesse M., and Patric H.. Hendershott, 1996, Bubbles in Metropolitan Housing Markets, *Journal of Housing Research* 7, 191–207.

Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

——— , 2003, Consumption-Based Asset Pricing, *Handbook of the Economics of Finance* , forthcoming.

——— , Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press: Princeton).

Campbell, John Y., and Robert J. Shiller, 1988, The Dividend-Price Ratio and Expectation of Future Dividends and Discount Factors, *Review of Financial Studies* 1, 195–228.

Capozza, Dennis R., Patric H. Hendershott, Charlotte Mack, and Christopher. Mayer, 2002, Determinats of Real Houses Price Dynamics, *NBER Working Papers* 5, 191–208.

DiPasquale, D., and W. Wheaton, 1994, Housing market dynamics and the future of housing prices, *Journal of Urban Economics* 35, 1–27.

Geltner, David M., and Norman G. Miller, 2000, *Commercial Real Estate Analysis and Investments* (South-Western Educational Publishing: Cincinatti Ohio).

Ghysels, E., P. Santa-Clara, and R. Valkanov, 2005, There is a risk-return tradeoff after all, *Journal of Financial Economics* 76, 509–548.

——— , 2006, Predicting volatility: getting the most out of return data sampled at different frequencies, *Journal of Econometrics* 131, 59–95.

Ghysels, E., A. Sinko, and R. Valkanov, 2006, MIDAS Regressions: Further Results and New Directions, *Ecnometric Reviews* , forthcoming.

Ghysels, E., and R. Valkanov, 2006, Linear Time Series Processes with Mixed Data Sampling and MIDAS Regression Models, UNC and UCSD working paper.

Glaeser, Edward L., Joseph Gyourko, and Raven Saks, 2004, Why is Manhattan So Expensive? Regulation and the Rise in House Prices, Harvard University Working Paper.

Karolyi, G., and A. Sanders, 1998, The Variation of Economic Risk Premiums in Real Estate Returns, *Journal of Real Estate Finance and Economics* 417, 245–262.

Lamont, Owen, and Jeremy Stein, 1999, Leverage and House Price Dynamics in U.S. Cities, *Rand Journal of Economics* 30, 498–514.

Lettau, Martin, and Sydney C. Ludvigson, 2005, Expected Returns and Expected Dividend Growth, *Journal of Financial Economics* 76, 583–626.

Ling, D., A. Naranjo, and M. Ryngaert, 2000, The Predictability of Equity REIT Returns: Time Variation and Economic Significance, *Journal of Real Estate Finance and Economics* 20, 117–136.

Liu, C., and J. Mei, 1992, The Predictability of Returns on Equity REITs and Their Co-Movement with Other Assets, *Journal of Real Estate Finance and Economics* 5, 401–418.

Malizia, Emil E., 1991, Forecasting Demand for Commercial Real Estate Based on the Economic Fundamentals of U.S. Metro Areas, *The Journal of Real Estate Literature* 6, 251–265.

Mayer, C., and C. Somerville, 2000, Residential construction: Using the urban growth model to estimate housing supply, *Journal of Urban Economics* 48, 85–109.

Miles, Mike, Rebel Cole, and David Guilkey, 1990, A Different Look at Commercial Real Estate Returns, *AREUEA Journal* 18, 403–430.

Nelling, E., and J. Gyourko, 2000, The Predictability Of Equity REIT Returns, *Journal of Real Estate Research* 16, 251–268.

Plazzi, Alberto, Walter Torous, and Rossen Valkanov, 2006, Expected Returns and the Expected Growth in Rents of Commercial Real Estate, UCLA and UCSD working paper.

Poterba, J., 1991, House price dynamics: The role of tax policy and demography, *Brookings Papers on Economic Activity* 96, 718–740.

Rosen, S., 1974, Hedonic prices and implicit markets: Product differentiation in pure competition, *Journal of Political Economy* 82, 34–55.

——— , and R. Topel, 1988, Housing investment in the United States, *Journal of Political Economy* 96, 718–740.

Sims, Christopher, 1972, Money, income, and causality, *American Economic Review* 62, 540–552.

Valkanov, Rossen, 2003, Long-Horizon Regressions: Theoretical Results and Applications, *Journal of Financial Economics* 68, 201–232.

Table 1: Forecasting Regression of Future Excess Returns on Log Cap Rate

The table reports the results from the OLS regression of future non-overlapping excess returns on a constant and log cap rate for each metropolitan area (MSA), as it appears in equation (3). The t-ratios, in parentheses, are Newey-West with 4 lags. The Table also reports the marginal economic significance of the cap rate on excess returns. The Table reports four entries: the first one ($ec_{\widehat{\beta}}$) corresponds to 2 times the standard error of the regressor times its coefficient, the second ($ec^{-}_{\widehat{\beta}}$) and the third ($ec^{+}_{\widehat{\beta}}$) correspond to the previous value where $\widehat{\beta}$ is replaced by a two standard deviations lower and upper bound, respectively. The last value ($ec_{\widehat{\beta}}/\sigma$) is the absolute value of $ec_{\widehat{\beta}}$ divided by the average return volatility. The sample is biannual observations for four property types from 1987:1 to 2002:2, for a total of $N = 64$ observations. The forecasting horizon is 1 year, or $k = 2$.

MSA	$\alpha_{i,2}$	$\beta_{i,2}$	R^2	$ec_{\widehat{\beta}}$	$ec^{-}_{\widehat{\beta}}$	$ec^{+}_{\widehat{\beta}}$	$ec_{\widehat{\beta}}/\sigma$
Atlanta	0.645	0.248	0.048	0.031	0.001	0.061	0.440
	(2.284)	(2.074)					
Baltimore	0.852	0.329	0.078	0.036	0.002	0.070	0.557
	(2.322)	(2.132)					
Boston	0.894	0.338	0.166	0.076	0.015	0.136	0.815
	(2.717)	(2.483)					
Charlotte	0.408	0.147	0.008	0.016	-0.030	0.062	0.175
	(0.818)	(0.696)					
Chicago	1.107	0.430	0.176	0.067	0.024	0.109	0.839
	(3.318)	(3.117)					
DallasFort Worth	1.501	0.615	0.261	0.082	0.048	0.116	1.021
	(5.038)	(4.871)					
Denver	0.574	0.212	0.019	0.023	-0.008	0.054	0.276
	(1.713)	(1.476)					
Houston	0.492	0.181	0.015	0.023	-0.017	0.064	0.245
	(1.368)	(1.157)					
Los Angeles	1.089	0.420	0.201	0.091	0.047	0.134	0.896
	(4.512)	(4.157)					
Minneapolis-St Paul	0.940	0.371	0.085	0.042	0.003	0.082	0.583
	(2.332)	(2.147)					

Table 1 (Cont'd): Forecasting Regression of Future Excess Returns on Log Cap Rate

MSA	$\alpha_{i,2}$	$\beta_{i,2}$	R^2	$ec_{\widehat{\beta}}$	$\underline{ec_{\widehat{\beta}}}$	$\overline{ec_{\widehat{\beta}}}$	$ec_{\widehat{\beta}}/\sigma$
Orange County	0.556	0.203	0.077	0.045	0.009	0.082	0.556
	(2.855)	(2.494)					
Orlando	0.997	0.388	0.073	0.041	0.006	0.075	0.541
	(2.572)	(2.369)					
Philadelphia	0.719	0.278	0.078	0.034	0.010	0.059	0.557
	(3.034)	(2.806)					
Phoenix	2.053	0.840	0.275	0.099	0.052	0.147	1.049
	(4.290)	(4.186)					
Riverside-San Bernardino	0.990	0.389	0.115	0.076	0.020	0.132	0.678
	(2.971)	(2.705)					
Sacramento	0.630	0.235	0.081	0.028	0.007	0.049	0.567
	(2.962)	(2.663)					
San Diego	1.214	0.472	0.169	0.078	0.028	0.128	0.822
	(3.319)	(3.099)					
San Francisco	0.437	0.153	0.053	0.054	-0.026	0.133	0.460
	(1.569)	(1.357)					
Seattle	0.544	0.197	0.025	0.023	-0.006	0.053	0.316
	(1.813)	(1.599)					
Tampa/St. Petersburg	1.226	0.489	0.116	0.050	0.013	0.087	0.681
	(2.879)	(2.709)					
Washington, DC	0.653	0.242	0.120	0.057	0.018	0.095	0.694
	(3.339)	(2.927)					

Table continued from previous page.

Table 2: Cap Rate Decomposition

The table reports the results from the OLS regression of log cap rate on a constant, orthogonalized economic variables (x_1 to x_4) and the rent growth component (y) for each metropolitan area (MSA), as it appears in equation (7). The t-ratios, in parentheses, are Newey-West with 4 lags. The sample is biannual observations for four property types from 1986:2 to 2000:2, for a total of $N = 64$ observations.

MSA	x_1	x_2	x_3	x_4	y	R^2
Atlanta	0.016	0.008	-0.030	0.045	-0.190	0.299
	(4.696)	(1.699)	(-5.721)	(2.086)	(-0.548)	
Baltimore	-0.017	0.012	0.001	-0.041	-0.155	0.230
	(-0.600)	(1.221)	(0.044)	(-5.507)	(-0.798)	
Boston	-0.013	0.007	-0.018	-0.056	-0.032	0.281
	(-0.465)	(0.491)	(-0.543)	(-3.001)	(-0.075)	
Charlotte	0.043	-0.008	-0.019	0.002	0.033	0.346
	(5.863)	(-3.847)	(-5.285)	(0.167)	(0.161)	
Chicago	0.041	-0.018	0.005	-0.019	0.486	0.325
	(2.698)	(-3.398)	(1.184)	(-1.592)	(3.737)	
DallasFort Worth	0.008	-0.018	0.002	0.015	0.076	0.113
	(0.503)	(-0.946)	(0.154)	(3.187)	(0.492)	
Denver	-0.012	-0.001	0.004	-0.013	0.420	0.122
	(-1.818)	(-0.120)	(0.638)	(-1.231)	(1.581)	
Houston	-0.005	0.005	0.000	0.016	0.038	0.123
	(-1.122)	(1.322)	(-0.010)	(3.887)	(0.227)	
Los Angeles	-0.128	-0.056	0.012	0.019	0.302	0.453
	(-10.963)	(-4.594)	(6.626)	(2.667)	(1.162)	
Minneapolis-St Paul	0.091	0.061	0.003	-0.012	0.064	0.252
	(4.399)	(4.628)	(0.455)	(-2.308)	(0.371)	

Table 2 (Cont'd): Cap Rate Decomposition

MSA	x_1	x_2	x_3	x_4	y	R^2
Orange County	-0.065	0.024	-0.015	0.018	-0.041	0.279
	(-4.741)	(3.110)	(-1.897)	(1.539)	(-0.148)	
Orlando	0.014	-0.038	0.035	-0.006	0.192	0.205
	(0.757)	(-4.760)	(3.582)	(-0.867)	(0.663)	
Philadelphia	-0.062	-0.003	-0.034	0.010	-0.222	0.422
	(-2.424)	(-0.533)	(-8.109)	(1.552)	(-1.309)	
Phoenix	0.031	0.062	0.035	0.019	-0.077	0.491
	(3.391)	(3.516)	(5.051)	(3.618)	(-0.452)	
Riverside-San Bernardino	-0.045	0.007	0.009	0.012	-0.180	0.673
	(-9.863)	(2.236)	(3.073)	(1.031)	(-0.848)	
Sacramento	-0.049	-0.008	-0.017	-0.014	0.207	0.211
	(-9.522)	(-3.067)	(-2.851)	(-1.171)	(0.885)	
San Diego	-0.045	0.018	0.022	-0.005	0.188	0.620
	(-7.194)	(4.038)	(2.039)	(-0.288)	(0.851)	
San Francisco	0.046	0.078	0.018	0.043	0.540	0.255
	(0.418)	(4.085)	(3.169)	(2.233)	(1.365)	
Seattle	-0.033	-0.005	-0.018	0.009	0.013	0.119
	(-3.444)	(-0.600)	(-1.171)	(2.164)	(0.094)	
Tampa/St. Petersburg	-0.035	-0.035	-0.024	-0.017	0.068	0.251
	(-1.120)	(-3.925)	(-1.998)	(-1.094)	(0.428)	
Washington, DC	0.060	0.049	-0.094	0.010	-0.820	0.351
	(1.908)	(1.528)	(-8.157)	(0.665)	(-1.618)	

Table continued from previous page.

Table 3: Decomposing the Predictability in Returns

The table reports the results from the OLS regression of future non-overlapping excess returns on a constant, orthogonalized economic variables (x_1 to x_4), rent growth component (y) and unexpected log cap rate component (v) for each metropolitan area (MSA), as specified in equation (8). The t-ratios, in parentheses, are Newey-West with 4 lags. The sample is biannual observations for four property types from 1987:1 to 2002:2, for a total of $N = 64$ observations. The forecasting horizon is 1 year.

MSA	x_1	x_2	x_3	x_4	y	v	R^2
Atlanta	-0.045	-0.040	-0.020	0.053	-0.072	0.049	0.333
	(-3.102)	(-2.228)	(-1.206)	(1.908)	(-0.325)	(0.448)	
Baltimore	0.023	-0.027	-0.052	0.011	-0.552	0.410	0.256
	(0.827)	(-2.150)	(-1.582)	(0.914)	(-1.549)	(2.228)	
Boston	0.023	-0.005	0.075	-0.061	-0.018	0.368	0.333
	(1.143)	(-0.399)	(3.621)	(-4.938)	(-0.047)	(3.119)	
Charlotte	-0.018	-0.027	-0.007	0.063	0.966	-0.102	0.226
	(-1.007)	(-5.716)	(-0.876)	(2.784)	(1.868)	(-0.435)	
Chicago	0.034	0.000	-0.003	-0.033	-0.020	0.410	0.270
	(1.049)	(-0.036)	(-0.309)	(-2.058)	(-0.099)	(3.249)	
DallasFort Worth	-0.029	-0.061	-0.008	0.027	0.484	0.522	0.421
	(-2.319)	(-3.043)	(-0.439)	(3.271)	(2.017)	(3.784)	
Denver	0.016	-0.006	0.034	-0.016	-0.105	0.344	0.393
	(2.135)	(-0.667)	(3.455)	(-1.275)	(-0.237)	(2.831)	
Houston	0.051	0.003	-0.036	0.031	0.822	0.310	0.461
	(4.074)	(0.311)	(-3.544)	(2.450)	(2.766)	(1.902)	
Los Angeles	-0.085	-0.070	-0.006	0.043	-0.028	0.182	0.391
	(-3.286)	(-2.372)	(-0.982)	(2.196)	(-0.130)	(1.380)	
Minneapolis-St Paul	0.093	0.092	0.033	-0.013	0.076	0.137	0.343
	(2.203)	(3.641)	(3.469)	(-1.223)	(0.411)	(0.963)	

Table 3 (Cont'd): Decomposing the Predictability in Returns

MSA	x_1	x_2	x_3	x_4	y	v	R^2
Orange County	-0.062	-0.012	-0.007	-0.041	0.105	0.113	0.255
	(-3.092)	(-1.163)	(-0.425)	(-2.703)	(0.443)	(1.405)	
Orlando	-0.059	-0.049	-0.024	0.002	0.350	0.431	0.268
	(-2.571)	(-2.713)	(-1.010)	(0.227)	(1.762)	(3.467)	
Philadelphia	-0.040	-0.012	-0.025	-0.018	0.009	0.175	0.160
	(-1.247)	(-1.821)	(-2.843)	(-1.637)	(0.036)	(1.615)	
Phoenix	0.029	0.059	0.013	0.057	-0.042	0.448	0.536
	(1.966)	(3.287)	(1.064)	(5.364)	(-0.156)	(3.234)	
Riverside-San Bernardino	-0.014	0.015	-0.001	-0.001	-0.263	0.120	0.189
	(-1.045)	(1.458)	(-0.116)	(-0.081)	(-0.744)	(0.566)	
Sacramento	-0.013	0.005	0.013	-0.032	0.332	0.175	0.163
	(-1.153)	(0.789)	(0.844)	(-1.721)	(1.400)	(1.792)	
San Diego	-0.035	0.002	0.027	-0.118	0.011	0.135	0.340
	(-1.632)	(0.170)	(1.409)	(-2.913)	(0.038)	(0.972)	
San Francisco	-0.086	0.018	0.012	0.001	0.223	0.057	0.163
	(-0.780)	(0.844)	(1.602)	(0.071)	(0.277)	(0.622)	
Seattle	-0.017	0.003	0.016	-0.022	-0.633	0.203	0.188
	(-0.788)	(0.321)	(1.102)	(-2.355)	(-5.742)	(1.556)	
Tampa/St. Petersburg	-0.039	-0.021	-0.030	0.012	0.285	0.334	0.205
	(-1.611)	(-0.794)	(-1.252)	(0.520)	(3.719)	(2.748)	
Washington, DC	0.071	0.095	-0.074	0.004	-0.191	0.101	0.249
	(1.765)	(2.681)	(-3.898)	(0.213)	(-0.590)	(0.915)	

Table continued from previous page.

Table 4: MIDAS Predictive Regressions

This Table shows the results for the MIDAS regression of equation (9). The parameters are estimated using non-linear least squares (NLS). The Table shows two different MIDAS regressions. The first one uses the exponential Almon parametrization, obtained by we estimating equations (9) and (10) jointly. The second one uses parameters from the Beta parameterization, obtained by estimating expressions (9) and (11) jointly. Standard errors are in parentheses below the estimates. The standard error of $1/\eta$ is obtained with the Delta method. The forecasting horizon in one year, or 260 trading days.

	φ	η	$1/\eta$	θ_1	θ_2	R^2
Almon	0.087	3.213	0.311	0.040	-0.001	0.354
	(0.023)	(1.293)	(0.103)	(0.009)	(0.001)	
Beta	0.079	3.056	0.327	1.602	5.498	0.325
	(0.021)	(1.182)	(0.993)	(0.548)	(1.305)	

Figure 1: Predictability and Its Provenance

This Figure shows the scatter plots of the cross-sectional coefficients of economic factors ($x1$ to $x4$), of the rent growth component y and of the orthogonalized "mispricing" variable v from Table 3 on the estimated coefficients of the cap rate $\widehat{\beta}_{i,2}$ from Table 1. for each of the plots, the figure also displays a linear fit along with the slope coefficient, its t-statistic and the R^2.

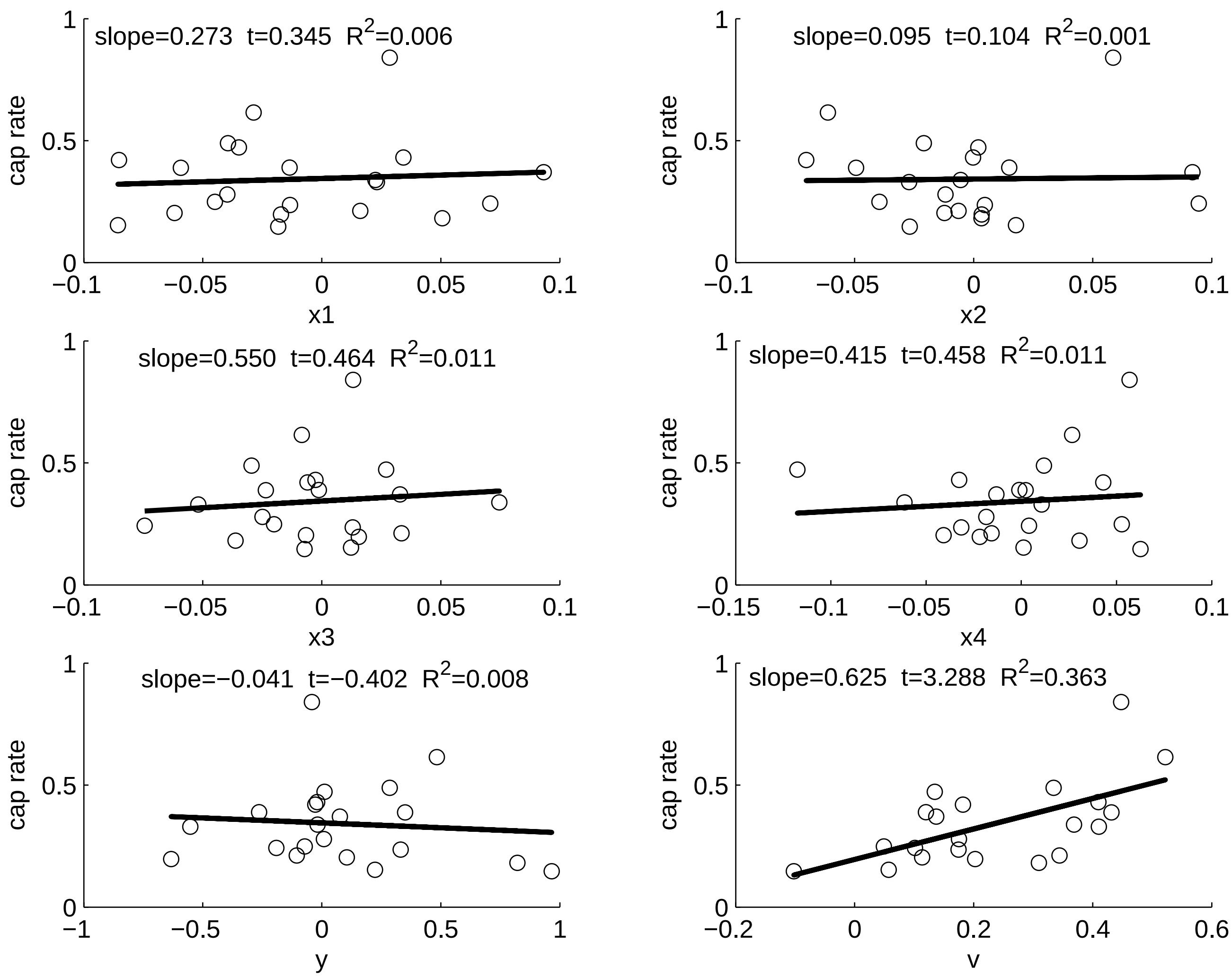
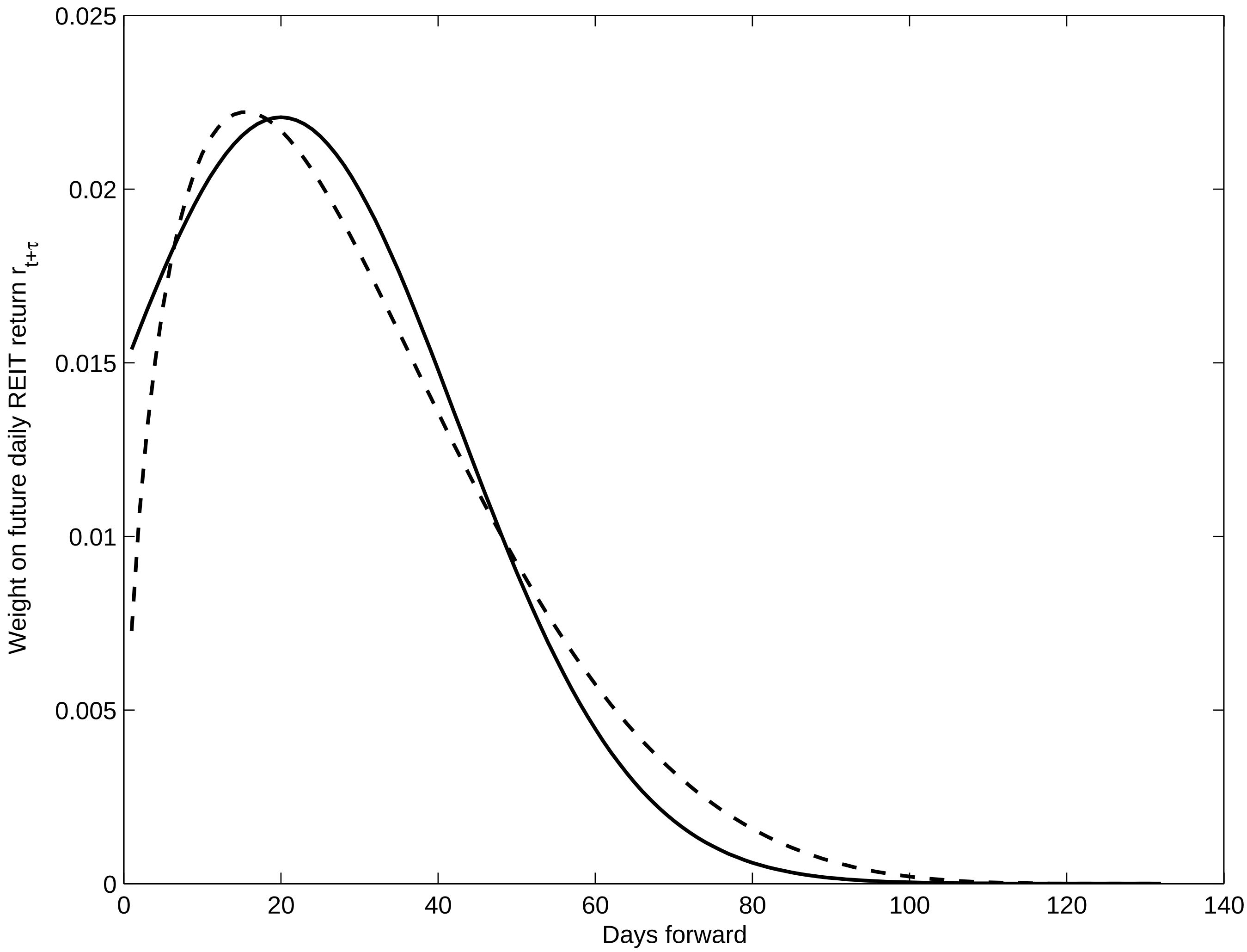


Figure 2: MIDAS Weights

The figure plots the weights that the MIDAS estimator (9) places on future daily REIT returns with exponential Almon and Beta weights. The weights are calculated by substituting the estimated values of θ_1, and θ_2 reported in Table 4 into the weight function (10). The horizon used in the estimation is 260 days. For expositional clarity, we plot only the first 130 weights. The remaining 130 weights are indistinguishable from zero and carry no information.



Appendix 1 - Metropolitan Areas

This table shows the list of MSAs and average values for real estate data used in the empirical section. The table reports the Metropolitan area (MSA) as well as averages of excess returns (denoted by r), rent growth (denoted by g) and annualized cap rates (denoted by cap), for apartments (superscript apt), industrial (superscript ind), retail (superscript rtl) and offices (superscript off). The sample is semi-annual observations of 21 areas between 1985:2 and 2002:2.

Metropolitan area	r^{apt}	r^{ind}	r^{rtl}	r^{off}	g^{apt}	g^{ind}	g^{rtl}	g^{off}	cap^{apt}	cap^{ind}	cap^{rtl}	cap^{off}
Atlanta	0.030	0.027	0.022	0.014	0.013	0.006	0.006	0.003	0.085	0.092	0.090	0.088
Baltimore	0.037	0.026	0.027	0.027	0.017	0.005	0.009	0.013	0.090	0.093	0.092	0.088
Boston	0.047	0.028	0.021	0.024	0.033	0.009	0.007	0.019	0.088	0.091	0.088	0.078
Charlotte	0.034	0.027	0.025	0.021	0.015	0.007	0.005	0.004	0.089	0.093	0.093	0.087
Chicago	0.036	0.032	0.021	0.016	0.015	0.012	0.005	0.001	0.088	0.092	0.088	0.083
DallasFort Worth	0.035	0.028	0.026	0.010	0.014	0.009	0.005	0.001	0.094	0.092	0.092	0.089
Denver	0.043	0.031	0.036	0.032	0.018	0.009	0.014	0.009	0.091	0.092	0.095	0.093
Houston	0.044	0.036	0.027	0.024	0.016	0.016	0.004	0.001	0.097	0.094	0.097	0.097
Los Angeles	0.043	0.020	0.024	0.011	0.022	0.003	0.005	0.005	0.085	0.088	0.091	0.079
Minneapolis-St Paul	0.038	0.029	0.027	0.020	0.019	0.009	0.005	0.007	0.088	0.096	0.095	0.089
Orange County	0.037	0.020	0.024	0.022	0.022	0.002	0.006	0.014	0.086	0.087	0.089	0.077
Orlando	0.033	0.039	0.029	0.023	0.014	0.018	0.011	0.007	0.089	0.094	0.091	0.088
Philadelphia	0.035	0.020	0.027	0.021	0.016	0.000	0.006	0.008	0.091	0.093	0.093	0.089
Phoenix	0.037	0.024	0.028	0.019	0.018	0.004	0.010	0.002	0.088	0.089	0.092	0.092
Riverside-San Bernardino	0.035	0.023	0.029	0.023	0.018	0.006	0.010	0.006	0.089	0.089	0.090	0.093
Sacramento	0.042	0.039	0.029	0.028	0.021	0.017	0.012	0.013	0.094	0.092	0.091	0.084
San Diego	0.038	0.033	0.027	0.024	0.020	0.013	0.009	0.008	0.084	0.090	0.090	0.083
San Francisco	0.040	0.025	0.020	0.019	0.025	0.010	0.006	0.020	0.085	0.089	0.087	0.074
Seattle	0.040	0.030	0.030	0.028	0.016	0.010	0.011	0.013	0.087	0.089	0.089	0.084
Tampa/St. Petersburg	0.030	0.032	0.031	0.014	0.010	0.013	0.012	-0.003	0.090	0.094	0.093	0.091
Washington, DC	0.040	0.027	0.023	0.030	0.019	0.006	0.009	0.017	0.087	0.094	0.087	0.078